

Examining the geologic link of arsenic contamination in groundwater in Orange County, NC

Ethan Dinwiddie

Honor's Thesis
The University of North Carolina, Chapel Hill
Department of Geological Sciences

April 10, 2017

Approved:

Abstract

In recent decades, elevated arsenic concentrations in groundwater have been observed at various locations around the world including the Carolina terrane North Carolina. Orange County lies in the middle of the Carolina terrane and previous studies have observed arsenic concentrations in private well water samples that exceed the EPA drinking water standard of 10 ppb. This study uses a large database of arsenic concentrations and other water quality measurements in private water wells collected by the North Carolina Department of Health and Human Services and whole rock arsenic analysis to examine a possible geologic connection to elevated arsenic concentrations in groundwater. Through the use of geostatistical modeling and multivariate statistics, a hydrogeologic connection between certain water quality variables and detectable arsenic is found.

Introduction

Arsenic is an extremely pervasive, naturally occurring, and potentially hazardous element that is found in air, soil, water, organisms, and rocks. Arsenic occurs in organic and inorganic forms, but inorganic arsenic compounds are far more toxic than organic arsenic compounds (Brown, K.G., Ross, G.L., 2002). Inorganic arsenic is mainly consumed by humans through drinking water (Kim, Miranda, Tootoo, Bradley, Gelfand, 2011) but can still be consumed if contaminated water is used for food preparation or irrigation (McCarty, Hanh, and Kim, 2011; WHO, 2016). Excessive and chronic lower level arsenic exposure is associated with numerous negative health effects including but not limited to death, lung and skin cancer, vascular and heart disease, skin problems, diabetes, and many more, including lesser ailments such as vomiting and diarrhea (Kim et al., 2011; Brown, Ross, 2002; McCarty et al. 2011). High

concentrations of naturally occurring arsenic in groundwater have been observed in locations globally, such as Southeast Asia, Chile, Argentina and the Western United States, which is of concern in regard to the potential health effects that people may experience in these places from high arsenic consumption (Nordstrom, 2002 Smedley, Kinniburgh, 2002; McCarty et al, 2011).

Currently in the United States, the concentration of arsenic in water is controlled by an EPA drinking water standard of 10 µg/L, which was reduced from 50 µg/L in 2001, although the Maximum Contaminant Level Goal (MCLG) is set at zero (USEPA, 2001). However, this standard is not enforced by the EPA for private wells in the United States so arsenic contamination in private wells could lead to chronic arsenic poisoning if the contamination is unnoticed. While extensive examination has been done on arsenic contamination in areas such as Southeast Asia, Wisconsin, and the northeastern U.S., recent private well water testing in North Carolina by the North Carolina Department of Health and Human Services (NCDHHS) has revealed high levels of arsenic in various locations throughout the state but mainly occurring in the Piedmont region. The main hypothesis for this study is that a natural connection exists between the bedrock and groundwater concentrations of arsenic. This study will test this hypothesis through geostatistical modeling, examining a direct geological connection through analysis of arsenic concentrations in bedrock, and multi-variate statistical analysis on water quality characteristics.

Arsenic geochemistry and release

Arsenite (As^{3+}) and arsenate (As^{5+}), in the form of oxyanions, are the dominant inorganic arsenic species found in groundwater (Maascheleyn, Delaune, Patrick, 1991; Welch, Westjohn, Helsel, Wanty, 2000) and is introduced to groundwater from anthropogenic and natural sources.

Anthropogenic sources can include mining, industrial, and agricultural activities while the natural source is dominantly from bedrock or aquifer release into groundwater (Hinkle, Polette, 1999; Smedley, Kinniburgh, 2001). Arsenic is naturally found in arsenic-bearing sulfide minerals, oxide minerals, and adsorbed onto metal oxyhydroxides. The main mechanisms for mobilization of this naturally sourced inorganic arsenic into groundwater supply are fairly well understood and mainly involve the oxidation of arsenic-bearing sulfide minerals, reduction of iron, manganese, and other metal oxyhydroxides with adsorbed arsenic, and replacement of adsorbed arsenic with other anions (Peters, 2008; Biwas, Hendry, Essilfie-Dughan, 2016; McCarty et al., 2011; Hinkle, Polette, 1999). The geologic setting as well as the chemical and mechanical properties of water, such as pH, reduction-oxidation potential, and water mobility in bedrock, are the dominant factors that control the mobilization of arsenic in groundwater in a given area (Hinkle, Polette, 1999). For example, in oxidizing conditions at near neutral pH, arsenate is the dominant species in water while in reducing and alkaline conditions, arsenite is the dominant species from desorption of arsenic-bearing metal oxyhydroxides (Welch et al, 2000; Hinkle, Polette, 1999). Overall, the environmental and geologic conditions of an area determine groundwater susceptibility and high arsenic concentrations in the bedrock or aquifer of an area do not always mean there will be high arsenic concentrations in the groundwater there (Smedley, Kinniburgh, 2002).

Geologic Setting

The geology of Orange County is associated with the Carolina terrane and specifically the Hyco Formation unit of the terrane. The Hyco Formation is comprised of Proterozoic age metaintrusive, metavolcanic, and metamorphosed volcanoclastic sedimentary rocks that date

from around 630 Ma to 613 Ma. Rocks in the Hyco Formation from oldest to youngest include felsic and dacitic lavas and tuffs, granodiorites, andesitic to basaltic lavas and tuffs, mixed epiclastic-pyroclastic rocks, gabbro, and then more granodiorite. Intruding into the Hyco Formation are the Neoproterozoic age East and West Farrington Plutons which are comprised of both intermediate and felsic plutonic rocks. Also intruding the Hyco Formation is the Neoproterozoic to Cambrian age Prospect Hill Pluton which is comprised of mainly granodiorite. Small scale dikes and intrusive bodies of Neoproterozoic to Mesozoic age spot the county until the deposition of Triassic age sandstone and siltstone that make up part of the Durham Basin (Bradley, Hanna, Gay, Stoddard, Bechtel, Phillips, 2016).

This study aims to examine the connection between elevated levels of arsenic and geologic units in Orange County by geostatistically analyzing an updated and more accurately geolocated database of arsenic concentrations from private wells and an updated geologic map of Orange County. Also by analyzing arsenic concentration in rock samples from each geologic unit in Orange County, a direct connection between arsenic concentrations in the rock and arsenic concentrations in the groundwater can possibly be established.

As part of the Carolina terrane, groundwater in Orange County has the potential to be at risk for elevated concentrations of arsenic. Approximately 40% of people in Orange County rely primarily on groundwater (Cunningham, Daniel, 2001), meaning thousands of people in Orange County have the potential to be unknowingly exposed to elevated arsenic concentrations through their drinking water if their wells are untested. This study will shed light on which areas are most affected by elevated groundwater arsenic levels and which rock units most likely increase arsenic concentrations, hopefully prompting well owners to begin monitoring their wells more regularly. This study hypothesizes that a spatial, geochemical, and hydrochemical connection between

geologic units and predicted arsenic concentrations in groundwater can be determined at the county level and that the arsenic in groundwater is mostly naturally sourced from dissolution of arsenic bearing minerals in the bedrock.

Previous work

Previous studies that have focused on geostatistically analyzing the arsenic distribution in North Carolina include Pippin (2005), Kim et al. (2011), and Sanders et al. (2012) to assess the risk it poses to public health and find the geologic connection to the elevated concentrations. Pippin (2005) was one of the first studies to use a large database to analyze the distribution and probability of arsenic concentrations in the state. This study used North Carolina DHHS data of 11,214 groundwater analyses and used ordinary point kriging with a linear drift algorithm to spatially analyze and create probability maps of groundwater arsenic concentrations exceedance of certain thresholds. The study found a connection to some underlying geologic units, most notably in the rock units associated with the Carolina terrane, which trends northeast from Union County to Person County. Additionally, Pippin's probability analysis found that Orange County is among the top counties that could host water supply wells that produce groundwater with at least 1 ppb of arsenic. A potential flaw with this study though is that the well point data was geolocated to address which could incorrectly place the well location and potentially place it in the wrong geologic unit. Overall though, Pippin's approach created a great generalized overview of connecting elevated arsenic in groundwater to geologic units in North Carolina.

Sanders, Messier, Shehee, Rudo, Serre, and Fry (2012) took a more public health risk approach and used 63856 well water samples to examine the statewide distribution and predict groundwater arsenic values in untested areas. Through the use of a four-class geocoding

algorithm and the use of a Bayesian Maximum Entropy geostatistical framework, more accurate predictions of well concentrations than classical kriging methods were made. Furthermore, a spatial correlation between the Carolina terrane and at-risk counties was observed, but a geologic connection was not specifically examined.

Kim et al (2011) looked at arsenic concentrations specifically in Orange County in connection to geologic units and depth of wells. In this study, Kim et al. (2011) used a standard batch geocoding protocol developed by the Children's Environmental Health Initiative at Duke University that geolocated the wells to tax parcel centroids instead of address which more accurately geolocates the wells. This study used only 471 data points obtained through the NC DHHS and an incomplete geologic map of Orange County with only the Chapel Hill, Hillsborough, and Efland quadrangles complete. They nonetheless found a connection between geologic units and well depth using spatial random effects modeling in a Bayesian computational framework. Overall, their findings were that wells close to transition zones and faults are more likely to contain detectable arsenic and that deeper wells and wells in welded tuffs and hydrothermal quartz bodies are associated with higher arsenic concentrations in groundwater (Kim et al., 2011).

Abraham (2009) expanded on the initial work done by Charles Pippin and directly investigated the geologic source-hypothesis between arsenic concentrations in groundwater and the bedrock source at a study site in Union County, which was shown to be a hotspot for arsenic contamination by Pippin (2005). The study drilled 2 wells into bedrock and 1 well into the regolith to monitor water quality and arsenic concentrations over time and monitored 23 nearby private wells for major and trace elements. In rock samples, Abraham (2009) was able to find a correlation between arsenic concentrations and Fe_2O_3 weight %, suggesting that arsenic is found

in Fe-bearing minerals and he found that precipitated sulfide and iron oxides exist along rock fractures. By taking a petrologic, geochemical, and hydrogeologic approach in the study, Abraham (2009) suggests that arsenic in the groundwater of the study site was naturally sourced and that the main mechanisms for release of arsenic comes from oxidation of iron-sulfide minerals and desorption of arsenic-bearing iron and manganese oxyhydroxides. Additionally, redox processes in the water related to pH heavily influence the release of arsenic from the bedrock (Abraham, 2009).

Data/methods

While the EPA does not monitor private wells, the NCDHHS (North Carolina Department of Health and Human Services) and local health departments began sampling private wells in 1999 (Pippin, 2005) under the statewide private well testing program (NCDEQ). Most of the data used, including a database of well water arsenic concentrations in Orange County, a database of water quality characteristics for wells across North Carolina, and a detailed bedrock map of Orange County, was provided by Phil Bradley of the North Carolina Geological Survey. The Orange County data contained 1335 arsenic concentration analyses of private wells that were geolocated to tax parcel centroids as described in Kim et al. (2011).

In order to assess the geologic correlation between arsenic concentrations in groundwater and the bedrock, the geologic units in Orange County were first simplified based on general rock type. These six general rock units were (i)felsic lavas and tuffs, (ii) felsic plutonic, (iii) intermediate/mafic plutonic, (iv) mafic lavas and tuffs, (v) Neoproterozoic epiclastics, and (vi) Triassic sedimentary.

The Orange County private well data contained 1335 arsenic concentration analyses of private wells that were geolocated to tax parcel centroids as described in Kim et al. (2011). The statewide well database contained 19443 samples but selecting samples only from Orange County reduced the sample size to 769 and these samples were then joined to the more accurately geocoded Orange County well data based on NC DHHS sample number. In these datasets, arsenic concentrations were in units of mg/L and sample concentrations that were below detection limit were marked “<0.001”, so a new field was generated to express these values as zero, since their actual concentration could not be assessed and a new field was created converting the concentration to $\mu\text{g/L}$ (ppb) for ease of examination. Other elemental concentrations and water quality parameters varied in units and detection level and any samples that had variables that were below their respective detection limits were replaced with zero. The well sample points were then spatially joined to the simplified geologic map. Ultimately, two datasets of private well water in Orange County were obtained. The first dataset contained 1335 samples of only location and arsenic concentrations and the second dataset contained 769 samples with location, arsenic concentrations, and many more water analyses variables.

Interpolation modeling

Kriging is an advanced geostatistical procedure that can be used to create surfaces of estimated values and probabilities based on a set of scattered points and their values. Previous studies such as Pippin (2005) and Yang (2009) have used indicator kriging methods to create probability maps of certain arsenic concentration ranges in groundwater and studies such as Kim et al. (2011) and Sanders et al. (2012) have used empirical Bayesian kriging methods to create probability and prediction maps. For this study, simple kriging modeling was used to create a

prediction map of arsenic concentrations for Orange County using arsenic concentration data from 1335 private wells to initially see if some sort of spatial connection existed between certain rock groups and higher or lower concentration predictions. This method was used because it allowed for transformation of the data to a normal distribution using normal score transformation. Normal score transformation works by ranking the dataset from lowest to highest values and matching these ranks to equivalent ranks from a normal distribution (ArcGIS Desktop Help 9.3). A normal distribution of the data significantly helps the accuracy of the kriging method because outliers can incorrectly influence kriging interpolations. The various parameters of the prediction model were then adjusted manually until the semivariogram model seemed to best fit the averaged semivariogram values. See “Supplements” (page) for images of the normal score transformation of the data and the semivariogram modeling and specifications. Geologic units that appeared to be in zones of high prediction values were added to the prediction map to show their possible spatial correlation. Statistics were calculated for arsenic concentration in each geologic unit using the “Summarize” tool.

Wholerock sample collection

In order to examine a direct connection between the concentration of arsenic in bedrock and in groundwater, bedrock samples of each simplified group were collected. This was done with the help of Philip Bradley and Brandon Peach of the North Carolina Geological Survey and an NCGS outcrop map of Orange County. 26 samples were collected in all from the six general units. Phil Bradley provided five of the 26 samples because the NCGS had them in stock from previous collection. Table 1 provides a summary of the number of samples and sample IDs from each unit.

Generalized unit	Number of samples	Sample IDs
Felsic lavas and tuffs	8	CH-04, CH-392, CH-1260*, CH-2311*, CD-15*, WX-304, WX-380, HL-3098*
Felsic plutonic	2	CH-418, CH-2488
Mafic lavas and tuffs	3	CH-374, HL-553*, HL-2294
Intermediate/mafic plutonic	3	CH-1069, CH-1196, CH-2207
Proterozoic meta-epiclastics	5	Epi-1, CH-367, CH-2093, CH-2266, CH-2267
Triassic sedimentary	5	Tr-1, Tr-2, Tr-3, CH-522, CH-523

Table 1. Summary of number of wholerock samples collected for each of the generalized units in Orange County. The sample ID corresponds to the quadrangle in which the samples were collected. CH=Chapel Hill quadrangle, CD=Caldwell quadrangle, HL-Hillsborough quadrangle, and WX=Whitecross quadrangle. Some of the sample IDs are named “Tr” or “Epi” because they were not located on the outcrop location map provided by the NCGS and were collected without the assistance of Phil Bradley and Brandon Peach. Samples directly provided by Phil Bradley were CH-2311, HL-553, HL-3098, CH-1260, and CD-15, noted by (*).

Phil Bradley provided previous whole rock analyses conducted by the NCGS on 49 samples to supplement the limited data from the collected samples. The detection limit for arsenic in these analyses was 3 ppm, which is high in comparison to the detection limit of the ICP-MS used in this study, so this data may not be as accurate and is treated with a certain degree of skepticism. Table 2 provides a summary of the number of samples located in each unit and the sample IDs.

Generalized unit	Number of samples	Sample IDs
Felsic lavas and tuffs	15	WX-28, WX-216, WX-899, WX-976, WX-1067, WX-1080, WX-1136, WX-1137, WX-1138, WX-4104, WX-4267, HL-277, HL-914, HL-2554, HL-2943
Felsic plutonic	9	WX-4095, WX-4362, EF-306, EF-599, EF-730, EF-2044, EF-2531, DF-3, DF-4
Mafic lavas and tuffs	13	EF-140, EF-368, EF-2001, EF-2142, EF-2187, EF-2207, HL-328,

		HL-419, HL-1177, HL-2233, HL-2418, HL-2643, HL-2752
Intermediate/mafic plutonic	6	WX-552, WX-4155, WX-4440, WX-4441, CH-533, DF-1
Proterozoic meta-epiclastics	6	DF-2, WX-1142, WX-1143, WX-1144, WX-1145, HL-3060
Triassic sedimentary	0	

Table 2. Summary of number of wholerock samples previously analyzed by the NCGS and matched to the six generalized rock units in Orange County.

Wholerock sample preparation/analysis

Sample preparation for whole rock analyses involved the use of a rock saw to carefully remove weathered parts of the samples so that a clean and unweathered piece was left. A Chipmunk Jaw Crusher was used to pulverize the unweathered pieces into smaller pieces, which were then powdered using a ball mill. The equipment used in the sample processing stage was meticulously cleaned using water and brushes so cross contamination between samples was minimal or non-existent. After powdering the samples, the Triassic sedimentary samples were put into an oven at 400° C for 3 hours on aluminum trays to incinerate possible organic material. This was done to eliminate organic arsenic that may have accompanied organic material in the sample because this study is focused on inorganic arsenic concentration in the samples. Based on the findings of Gray, Watts, and Overcamp (2001), 400° C was used because it is less than the temperature of 522° C at which inorganic arsenic volatilization was observed but sufficiently high enough to burn off organic material. Some inorganic arsenic could have volatilized during this burning process though which is why both unheated and heated Triassic sedimentary samples were analyzed in case this happened.

Approximately 50 milligrams of powder was weighed out per sample based on concentration calculations that assumed at least 0.1 ppm of arsenic in each of the samples. In addition to the collected samples, 4 standard samples were selected to compare the arsenic

concentration results. These standards were USGS standards BHVO-2, GSP-2, and SBC-1, and a kaolinite sample from the Galápagos. Dissolution of the rock powder was done using a step acid digestion method of concentrated hydrofluoric and concentrated nitric acid to initially dissolve the rocks and then concentrated hydrochloric acid to dissolve the remaining fluoride crystals. In some samples, such as the USGS SBC-1 shale standard, aqua regia (1 part concentrated nitric acid: 3 parts concentrated hydrochloric acid) was used to dissolve the sample if the hydrofluoric and nitric acid mix did not dissolve it completely. During all stages of the acid digestion, the samples were sealed in beakers and placed on a hot plate at approximately 140° C and were dried down in between steps at approximately 80°C. After residual material and fluoride crystals were dissolved and the samples were dried down, the samples were prepped for analysis on the ICP-MS by diluting them in 5 ml of 2 v/v % nitric acid and then again by taking 1 ml of this solution and diluting it in another 4 ml of 2 v/v % nitric acid. This was done to reduce the matrix percentage and avoid interferences on the ICP-MS. Helium gas was also used in addition to the carrying argon gas to reduce plasma- and matrix-based polyatomic interferences in both iron and arsenic analysis. Standard calibration solutions were made for both arsenic and iron after an initial calibration run. The arsenic calibration solutions were created from the dilution of an initial standard of 10 ppm and the iron calibration solutions were created by dilution of an initial standard of 997 ppm. Table 3 shows the calculated concentrations of each of the diluted standards after using a scale to initially weigh out the volume of standard used for dilution. Arsenic and iron concentrations analysis of the samples and standards was done using an Agilent 7900 Q-ICP-MS on hot plasma mode. Final concentration amounts in ppm were calculated accounting for dilution factor and initial sample mass.

Standard number	Arsenic (ppb)	Iron (ppb)
-----------------	---------------	------------

1	Blank (only 2% nitric acid)	Blank (only 2% nitric acid)
2	0.05463	4.174
3	0.5463	41.74
4	5.463	417.4
5	54.63	4174

Table 3. Calculated concentrations of standards after dilution to make calibration curve. The different levels of dilution were done by taking 1 ml of the standard above it and diluting it 9 ml of 2% nitric acid. The initial volume was calculated out using a scale and was 0.4189 ml for the 997 ppm iron standard and 0.5466 ml for the 10 ppm arsenic standard.

Results

Modeling and ICP-MS analysis

Initial mapping arsenic concentrations for each private well point from the 1335 samples in Orange County (Figure 1) appears to show that some sort of clustering or directionality exists in similar concentrations of arsenic. Generally, it appears that most of the wells in plutonic bodies have arsenic concentrations below detection limit (< 1 ppb) and that most of the wells with detected arsenic reside in the “felsic lavas and tuffs” and “Neoproterozoic epiclastics” units. Through arsenic concentration prediction mapping using simple kriging this relationship becomes more obvious. Figure 2 shows that there appears to be a good spatial overlap of the higher arsenic concentration prediction contours and the “Neoproterozoic epiclastics” and that the general direction of arsenic contamination trends northeast with the Carolina terrane. However, there are pockets of higher arsenic predictions in the “Neoproterozoic epiclastics” and another minor northwest direction to arsenic predictions. This could either be explained by error produced from the simple kriging modeling or that other variables besides bedrock are involved

with arsenic concentrations in well water and are affecting its release from minerals in specific locations within the “Neoproterozoic epiclastics” unit.

Well water averages

Group	Counts per group	Avg. As (ppb)
Felsic lavas and tuffs	644	0.723603
Felsic plutonic	243	0.283951
Intermediate/mafic plutonic	96	0.760416
Mafic lavas and tuffs	37	1.72973
Proterozoic epiclastics	307	1.247557
Triassic sedimentary	8	1.875

Table 4. Average arsenic concentrations in well water from each generalized rock grouping.

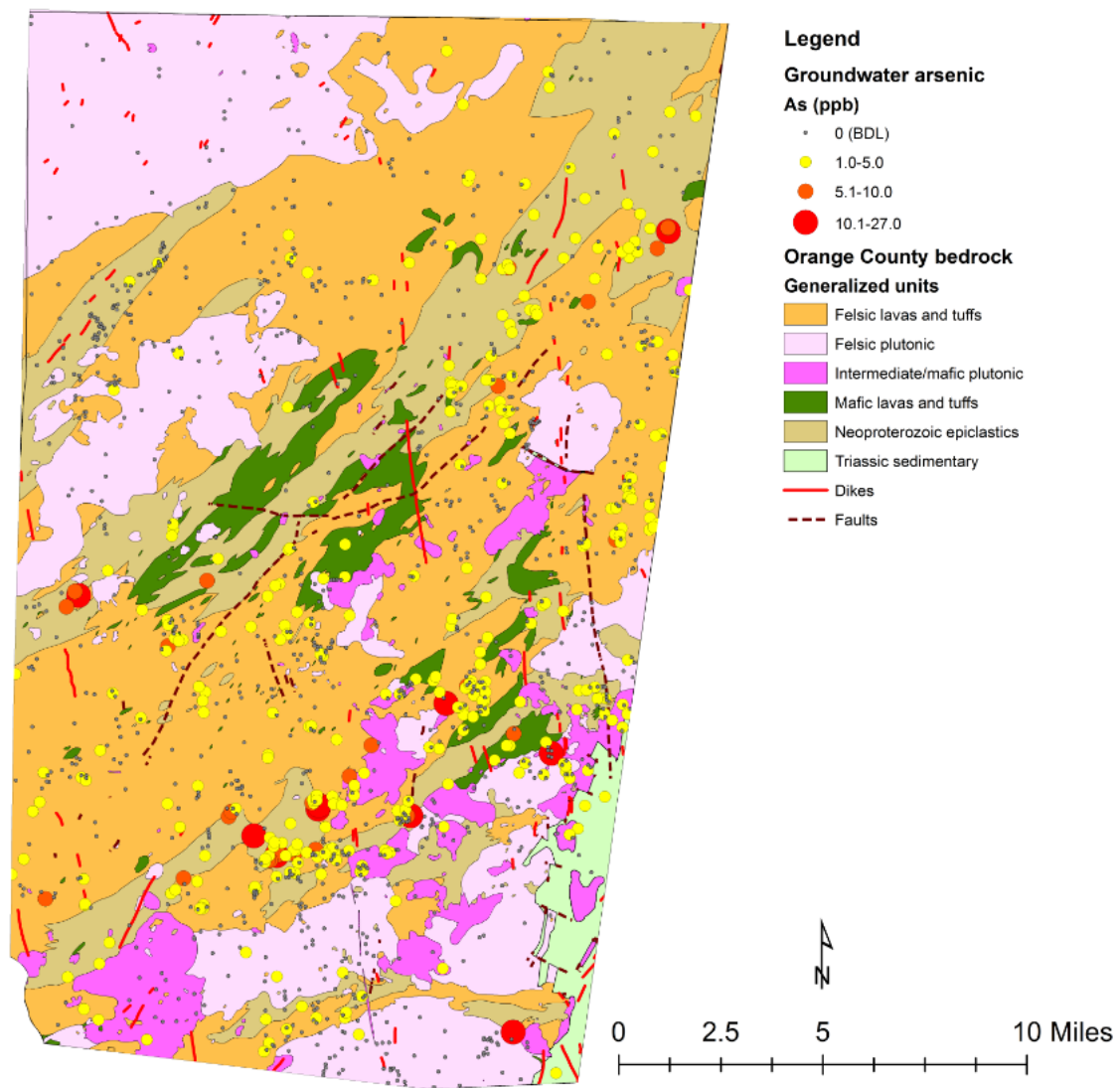


Figure 1. Generalized geologic map of Orange County with the well data of arsenic concentrations from groundwater sample data provided by Phil Bradley and obtained through the NC DHHS.

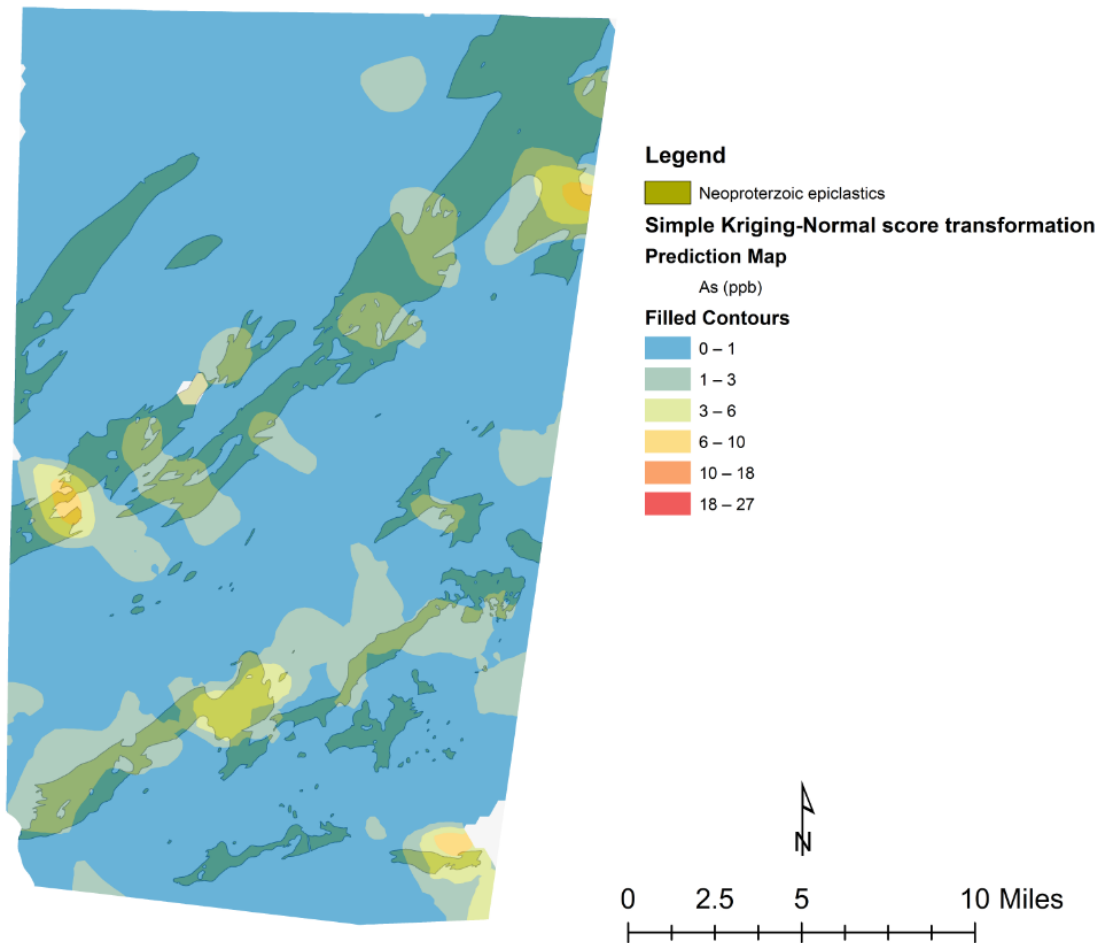


Figure 2. Neoproterozoic epiclastics unit underlying the arsenic concentration prediction map created with the simple kriging method.

This seemingly good spatial connection between the “Neoproterozoic epiclastics” and high predicted arsenic as well as high average arsenic concentrations in groundwater in the “Mafic lavas and tuffs” and “Neoproterozoic epiclastics” could possibly mean that a relationship between average arsenic content in rock and average arsenic content in groundwater exists. By analyzing arsenic concentrations of whole rock samples from each rock type, mean arsenic concentrations for each rock type were calculated. Previous whole rock analysis by the NCGS of 49 samples was also examined but kept separate from this study’s whole rock arsenic analysis to see if there was a difference. Overall, there appears to be much higher variation in mean arsenic

across rock types from the previous NCGS analysis than this study's analysis. This distinct difference could be from a number of reasons and perhaps because the NCGS analyzed samples more widely distributed across Orange County they were able to collect more representative arsenic concentrations of each rock type. Percentage error was calculated for the USGS standards that were analyzed versus known concentrations in those standards. For the SBC-1 shale standard, there was 8.4% error and for the BHVO-2 basalt standard, there was 54% error.

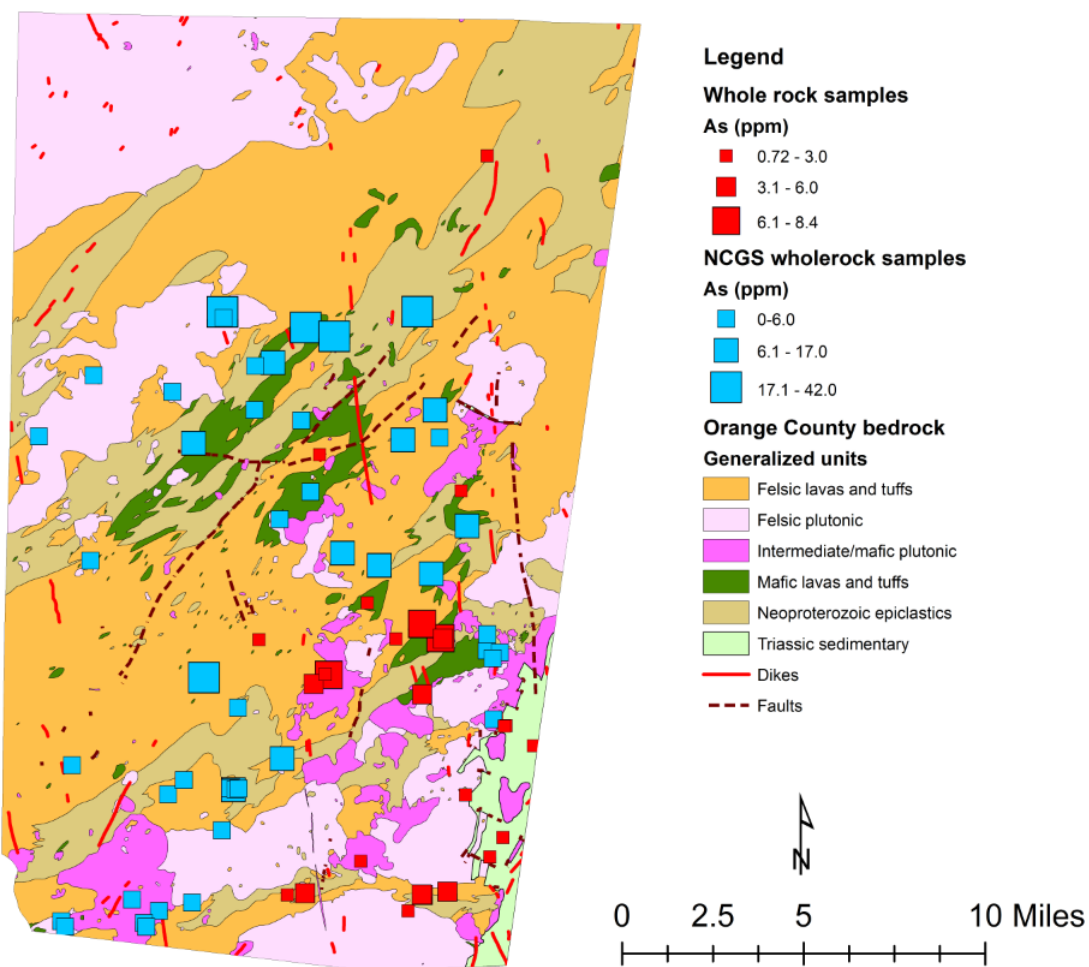


Figure 3. Generalized map of Orange County showing sampling locations from this study and previous NCGS wholerock analysis.

Examining the relationship between iron and arsenic concentrations in the rock shows some slight positive correlation between the two variables in each generalized rock unit meaning that arsenic is likely associated with iron in the mineralogy. Higher iron in the “intermediate/mafic plutonic” and “mafic lavas and tuffs” shows that iron is associated with other elements besides arsenic in the mineralogy, which makes sense because they are mafic rocks and tend to have more iron-bearing minerals. This relationship is not as clear or is much weaker with the NCGS whole rock analysis. Most of these relationships are fairly weak and given the number of data points there isn’t much more that can be said besides there is a general positive relationship between the two variables.

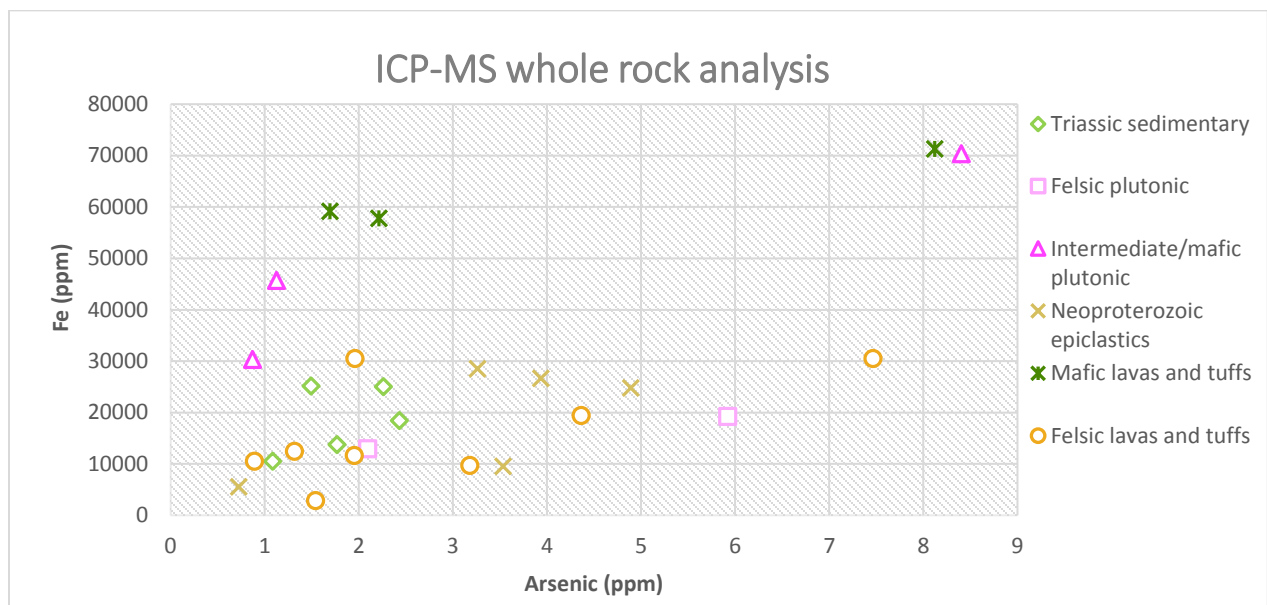


Figure 4. Scatterplot showing arsenic and iron concentrations of each whole rock sample analysis.

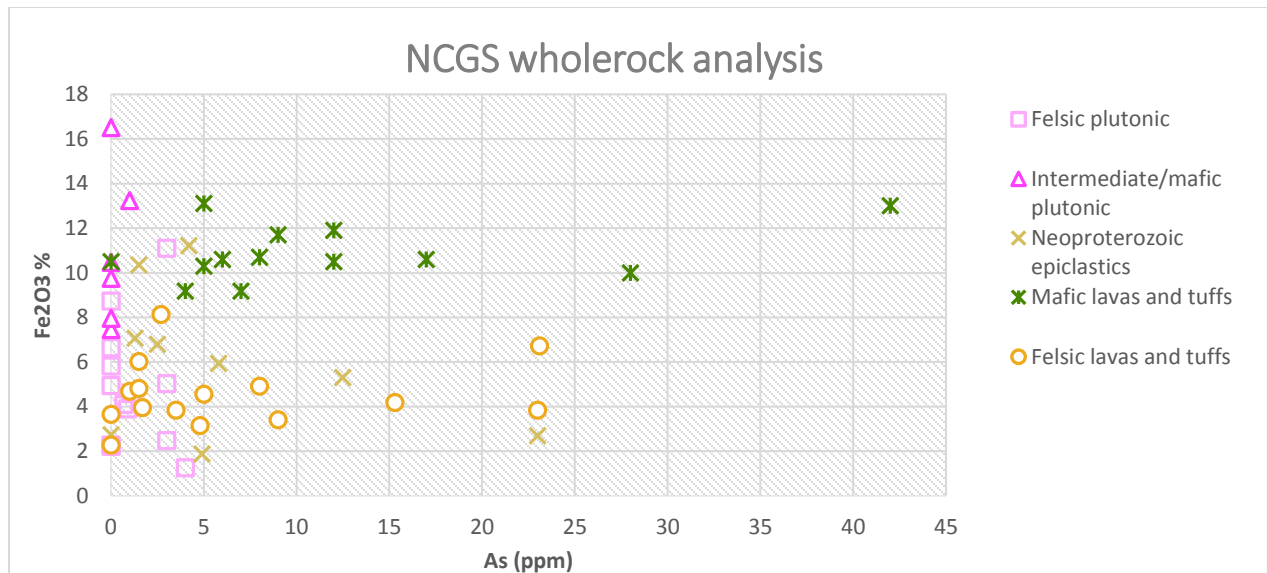


Figure 5. Scatterplot showing arsenic concentrations and iron oxide percentage for the NCGS whole rock analyses.

Examining relationships between average groundwater arsenic concentration in each rock unit and the average arsenic whole rock content for each rock yields a fairly strong relationship with the NCGS samples and a very weak relationship with the averages from this study's samples. These plots seem to provide evidence that there isn't a direct connection between bedrock concentrations of arsenic and arsenic concentrations in groundwater.

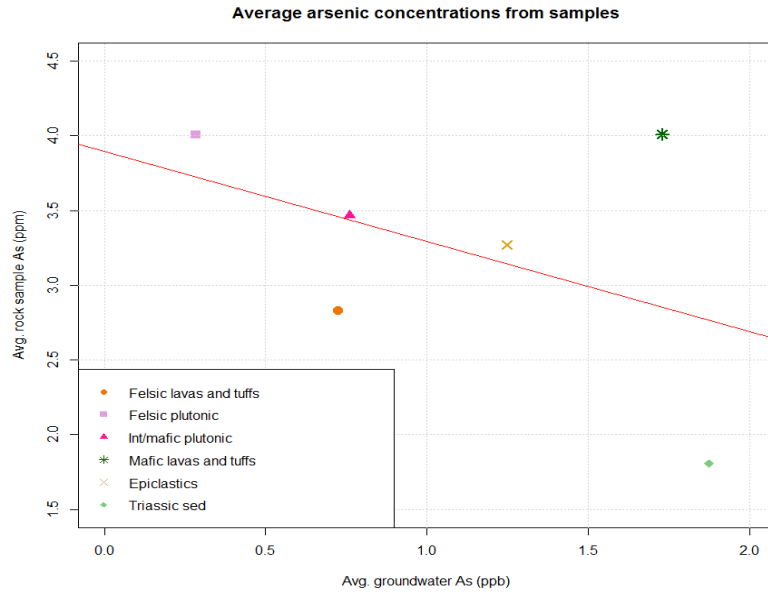


Figure 6. Average arsenic concentrations in water and rock, showing weak, negative relationship. $y = -0.6305x + 3.8796$, $r^2 = 0.205$

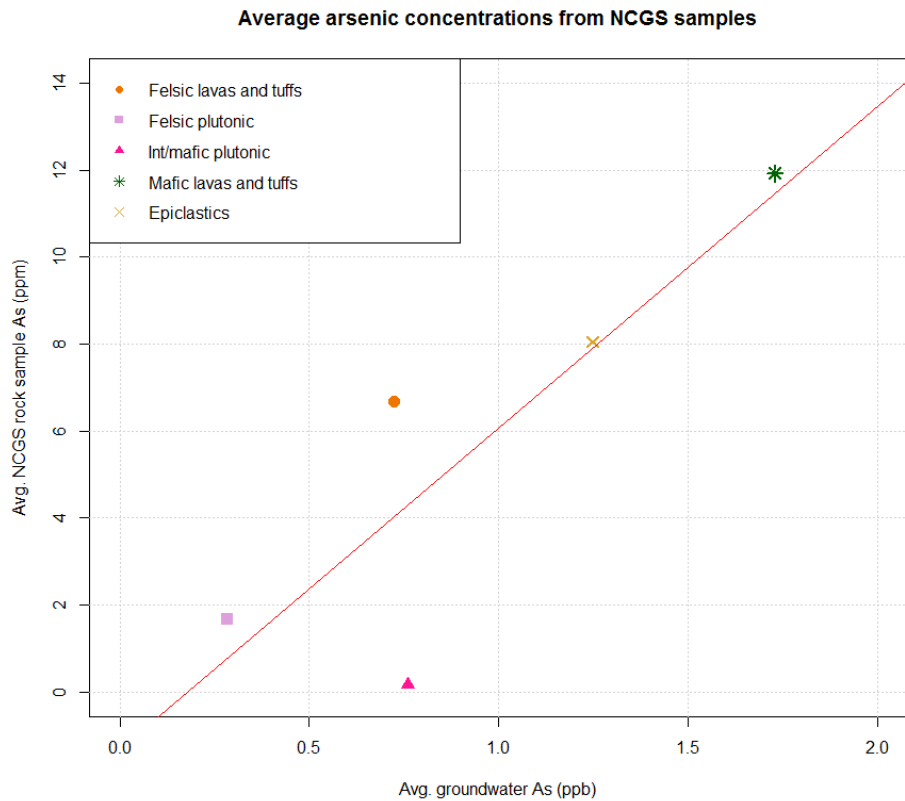


Figure 7. Average arsenic in groundwater and NCGS rock analysis, showing strong, positive correlation. $Y = 7.387x - 1.315$, $r^2 = 0.73$

Possible increases in proximity to plutons

Pluton emplacement in Orange County occurred in a few time periods: 1) Neoproterozoic (ca. 630 Ma); 2) Neoproterozoic (ca. 613-614 Ma); 3) Neoproterozoic (ca. 579 Ma); and 4) during the Cambrian/late Neoproterozoic (ca. 546 Ma). These emplacements occurred after the initial deposition of the felsic lavas and tuffs that occurred around ca. 629-633 Ma. It is possible and likely that these emplacements caused hydrothermal or other alterations to the surrounding and overlaying rocks and possibly increased the arsenic content at these boundaries. By using the near tool in ArcGIS, sample locations located within 500 meters or less to the plutons were selected and put in a separate shapefile. The rest of the sample locations were also put into a separate shapefile for ease of analysis. A two-tailed Mann-Whitney-Wilcoxon nonparametric test was used to test differences between the arsenic concentrations in wells closer than 500 meters to a pluton contact and arsenic concentrations in wells farther than 500 meters away from a pluton contact. This test did prove significant, but not in the way hypothesized because the average arsenic is higher away from the 500 meter buffer around the plutonic bodies.

	Average As (<500 m)	Average As (>500 m)	p-value
As (ppb)	0.69	0.88	9.487e-05 < 0.05=different

Table 5. Means of arsenic from each group with the p-value showing significance at the 95% confidence interval.

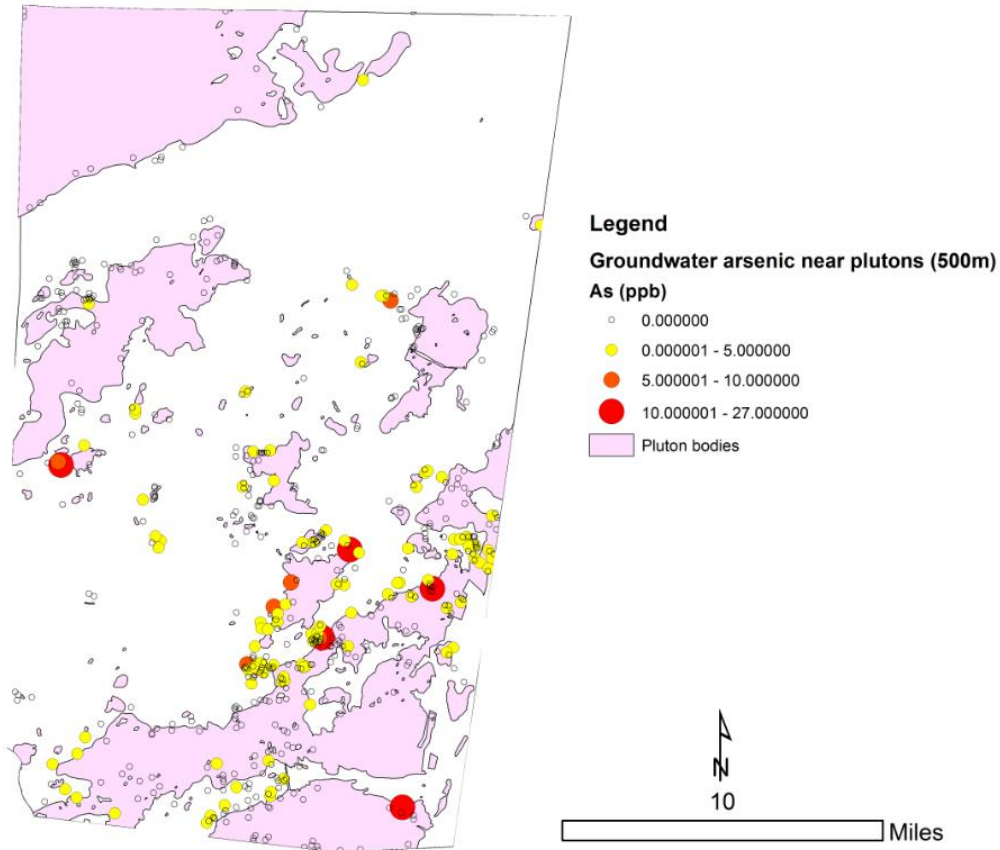


Figure 8. Map of Orange Counties with only the pluton bodies and proximal (within 500 m) well sample points shown.

Multivariate analysis

From the ICP-MS analyses, it seems that arsenic concentrations in the bedrock do not directly relate to arsenic concentrations in the groundwater meaning that some other factor must be influencing arsenic concentrations in the groundwater. The other dataset of 769 well water samples contained data for 12 other water quality parameters besides arsenic, including dissolved solids concentrations (such as Mg, Ca, Fe), pH, hardness, and alkalinity. Combining this data with distances of each well sample point to the nearest rock unit for each of the six rock types and the average arsenic concentrations from each rock type from my sample analysis, the NCGS

sample analysis, and the mean of the two analyses. This gave me a matrix of 769x21 with 21 variables to examine for possible correlations to detected arsenic in groundwater. To do this I first used hierarchical clustering analysis to get a quick idea of the groupings of the variables. The grouping from the hierarchical clustering analysis that contained arsenic concentration in wells was also grouped with Mg, Ca, hardness, F, alkalinity, and pH, meaning that they are all related in some way.

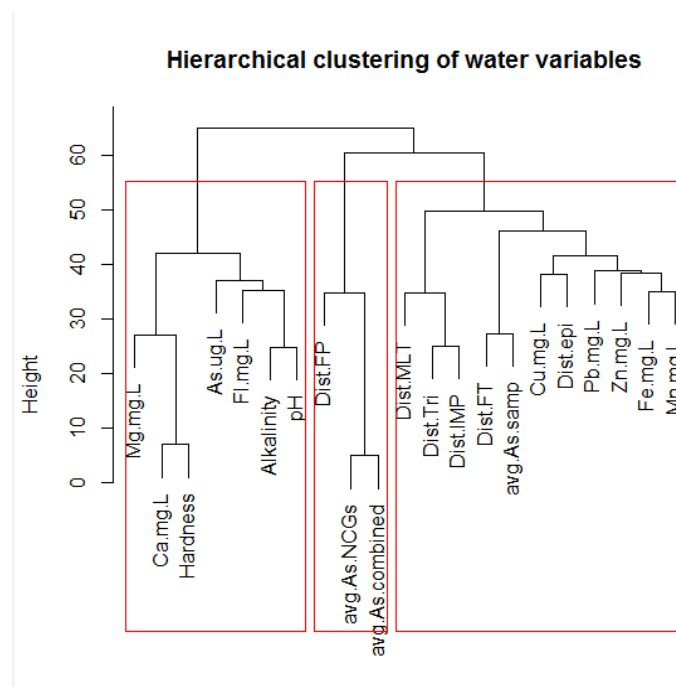


Figure 9. Hierarchical clustering dendrogram showing closely related variables in three groups that are boxed in red.

Based on this new group, Q-mode principal component analysis was performed to look at patterns of similarity in the samples over the variables in the new group. Q-mode PCA does this by explaining the variance in the matrix by principal components. Bi-plots of the Q-mode analysis were created to plot variables and samples and to assess other groupings within the samples such as rock grouping and arsenic detect in the sample. The sample points in the bi-plot

are plotted by their scores in principal component 1 and 2 and the variable vectors are plotted by their loadings in principal component 1 and 2. Sample points that are close together correspond to observations that have similar scores on the components displayed in the plot and if the points fit the variables well then they will plot closer to them. Likewise, variable vectors that point in the same general direction have positive correlation and negative correlation if they point opposite to each other. In Figure 10, the variables Mg, Ca, and hardness all point in the same direction and the variables As, pH, F, and alkalinity all point in the same direction indicating that these groups are correlated to one another. Also in Figure 10, the ellipses showing rock types do not seem to create distinguishable groups based on the principal components and thus not much variability in these groups is explained by these variables. In Figure 11, a somewhat distinct grouping is observed between wells with detected and no-detected arsenic. Although there is some overlap in the groups the differences in these groups should be examined in relation to these variables, especially pH, As, F, and alkalinity because the greatest shift in the groups is seen in those variables directions.

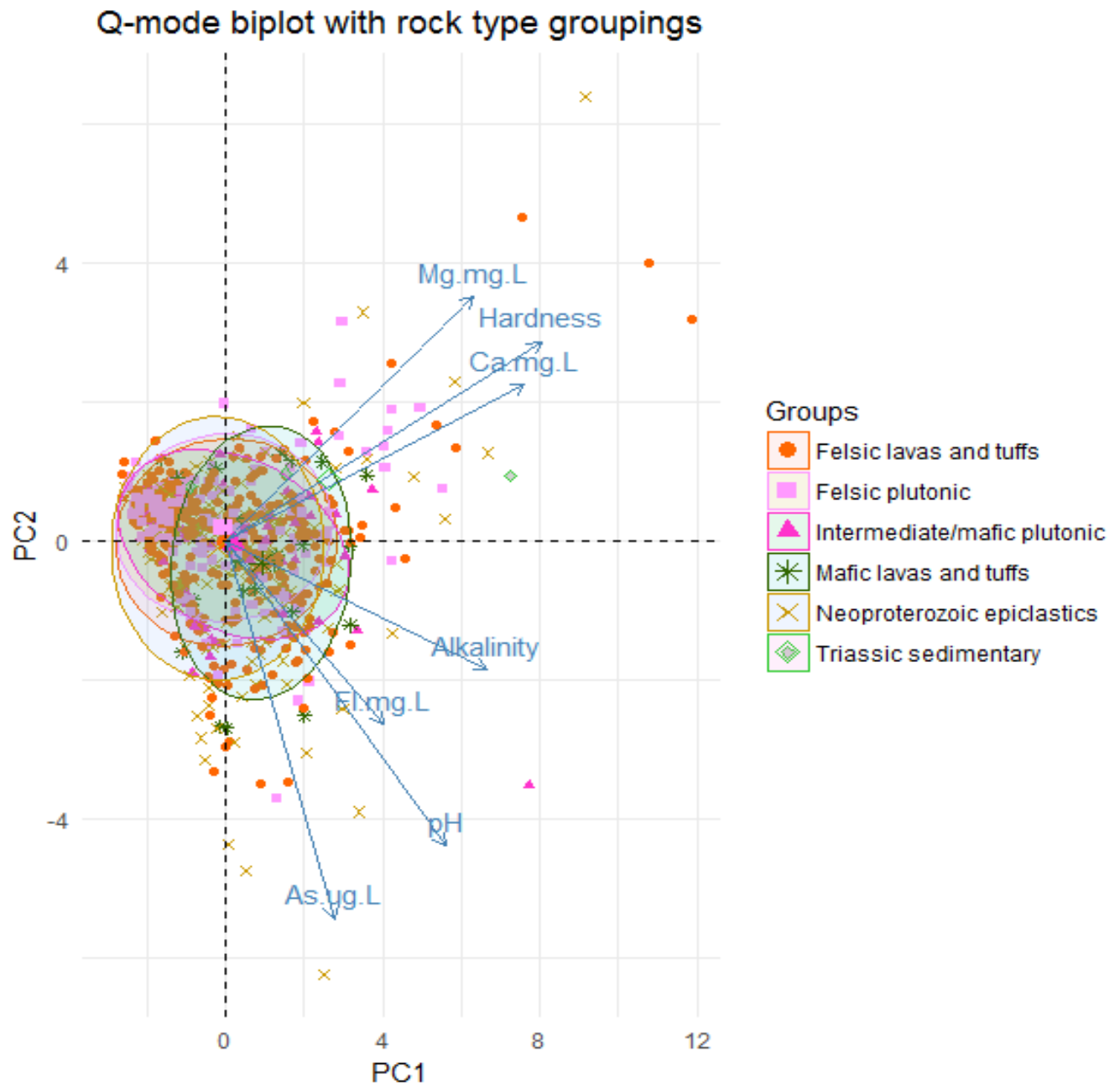


Figure 10. Q-mode PCA analysis biplot of the 1st two principal components. Ellipses were drawn around the points from each rock group encompassing 68% of the points in each group with each ellipse.

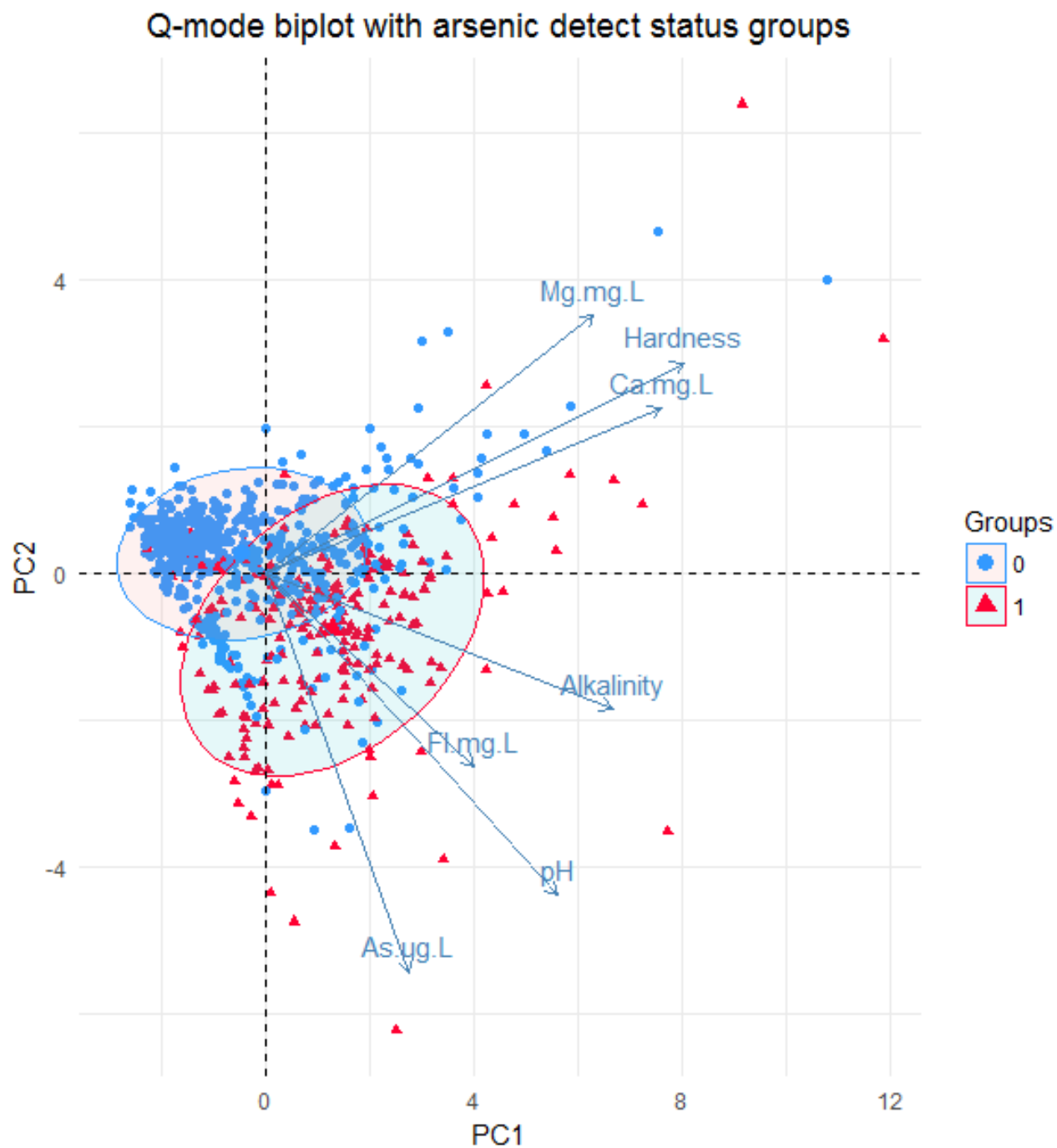


Figure 11. Q-mode PCA analysis biplot of the 1st two principal components. Ellipses were drawn around the points from each detect group encompassing 68% of the points in each group with each ellipse. Group=0=no detect, Group=1=detect.

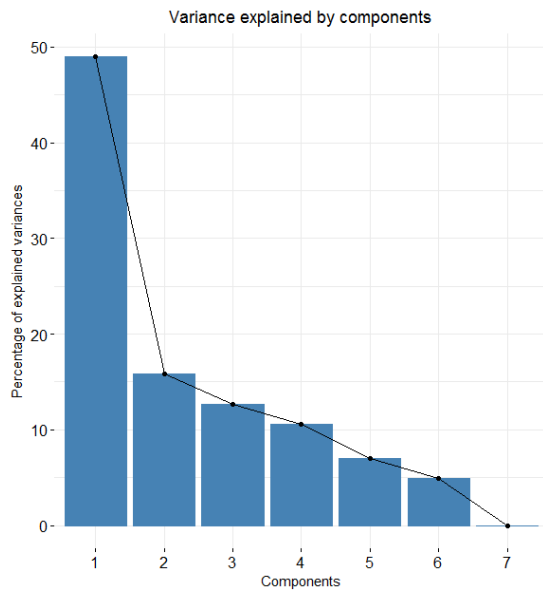


Figure 12. Histogram showing the percentage of variance in the dataset explained by each principal component.

To test differences between the non-detect and detect groups, the dataset was first divided into these two groups and then the variables pH, F, alkalinity, Mg, Ca, and hardness were selected out of each group. Shapiro-Wilk tests, at the 95% confidence interval, were performed on each of the variables in both groups to test for normality. Most of the tests came back rejecting the null hypothesis that the samples are normally distribution meaning the samples do not follow a normal distribution. Since most of the samples followed a non-normal distribution, the Mann-Whitney-Wilcoxon test was used to test differences between variables between the non-detect and detect groups because it is a nonparametric test. The Mann-Whitney-Wilcoxon tests for the paired variables from the two groups all showed that there is significant difference between the groups at the 95% confidence interval. Table 6 provides a summary of the means of each variable for each group and if there was significant difference.

	Mean value for detect samples	Mean value for non-detect samples	P-value
pH	7.4	6.9	$2.2\text{e-}16 < 0.05$ =difference
Alkalinity	139	88	$2.2\text{e-}16 < 0.05$ =difference
Mg (mg/L)	7.1	5.1	$5.188\text{e-}06 < 0.05$ =difference
Ca (mg/L)	39.5	19.6	$2.2\text{e-}16 < 0.05$ =difference
Fl (mg/L)	0.09	0.05	$2.615\text{e-}05 < 0.05$ =difference
Hardness	128	70.3	$2.2\text{e-}16 < 0.05$ =difference

Table 6. Means for each variable tested in each group and p-value from Mann-Whitney-Wilcoxon test.

Discussion

This study aimed to explain the elevated arsenic concentrations observed in some private wells in Orange County as naturally sourced. The results from spatial modeling and multivariate statistics seem to indicate that there is indeed a natural connection.

Initial and albeit, rudimentary, spatial modeling indicates that a spatial connection exists between the “Neoproterozoic epiclastics” and areas of higher predicted arsenic but this strong connection is not seen everywhere in the unit. Therefore, examining water quality measurements from roughly half of the wells using cluster and principal component analysis seem to indicate that arsenic is associated with pH, alkalinity, hardness, Mg, F, and Ca. All of these variables are statistically different between detect and non-detect arsenic wells using a non-parametric test at the 95% confidence interval and show higher means in the arsenic detected wells. These connections make sense from a geochemical perspective because the dominant factors that influence arsenic speciation and mobilization in water are pH, redox potential, alkalinity, and total dissolved solids. Increased pH in groundwater and alkalinity can lead to desorption of

arsenic from metal oxyhydroxides (Welch et al, 2000; Hinkle, Polette, 1999). Selective desorption through anion exchange between arsenic-bearing minerals and oxyanions such as F and bicarbonate has been observed to mobilize arsenic also (Casentini, Pettine, Millero, 2010). These are just two release mechanisms that could potentially explain the detected arsenic in wells that had increased pH, alkalinity, F, and hardness. Potential oxidation of arsenic sulfide minerals from well installation and drawdown of water table could also mobilize arsenic in water very locally (Abraham, 2009). Furthermore, high alkalinity has been associated with felsic volcanic rocks (Welch et al., 2000) and the geology of Orange County is dominated by felsic volcanic rocks representing a potential geologic link that went undetected during this analysis.

An assessment on the direct connection between arsenic concentrations in rocks and water cannot be made at this time because of the conflicting sample data from this study's whole rock analysis and previous NCGS analysis. While the NCGS sample data indicates a strong positive correlation between arsenic concentrations in rocks and groundwater, the sample data from this study does not indicate such a strong correlation. Additionally, whole rock analysis indicates that arsenic is associated with iron in rocks in Orange County, although there is no evidence to show that this association is from arsenic-bearing iron oxides/oxyhydroxides or associated with arsenopyrite.

Conclusion

This study found some potential mechanisms for natural arsenic release into groundwater involving association of arsenic in wells and higher pH, alkalinity, F, and hardness and a spatial relationship between the "Neoproterzoic epiclastics" unit and higher predicted arsenic from simple kriging modeling. This relationship could exist from either poor modeling of arsenic

prediction or some specific hydrologic characteristic(s) of the unit that have not yet been measured. It was clear from the PCA analysis that distances from each rock type, average whole rock arsenic concentration in each unit, and arsenic in groundwater is not related. It is also unlikely that arsenic is anthropogenically derived given the large areal distribution of elevated arsenic and the ready adsorption of arsenic onto clays and iron oxides in the soil and regolith. Overall, gathering more water quality data from private wells, including well depth, and analyzing arsenic speciation in rocks and water will enormously help our understanding of the dominant unit of origin for arsenic release and processes governing release in Orange County, NC.

Acknowledgements

I would like to thoroughly thank:

Dr. Xiao-Ming Liu, Dr. Clément Bataille, and the rest of the PMS lab members for their guidance, patience, and contributions involving sample and statistical analysis.

Phil Bradley and Brandon Peach of the NCGS for their geological expertise of the area and data.

Dr. Coleman, Dr. Glazner, and Dr. Mills and the rest of the isotope geochemistry/geochronology lab members for their assistance in dissolving my whole rock samples.

Citations

About Private Water Wells, 2016, EPA, <https://www.epa.gov/privatewells/about-private-water-wells> (access April 2017).

Abraham, J., 2009, A hydrogeological assessment of arsenic in the Carolina terrane of Union County, North Carolina: North Carolina Division of Water Quality, Groundwater Circular #2009-01, p. 1-40.

ArcGIS Desktop Help 9.3 - Normal score transformation, [http://webhelp.esri.com/arcgisdesktop/9.3/index.cfm?TopicName=Normal score transformation](http://webhelp.esri.com/arcgisdesktop/9.3/index.cfm?TopicName=Normal%20score%20transformation) (accessed April 2017).

Bradley, P.J., Hanna, H.D., Gay, N.K., Stoddard, E.F., Bechtel, R., Phillips, C.M., and Fuemmeler, S. J., 2016, Geologic map of Orange County, North Carolina: North Carolina Geological Survey Open-file Report 2016-05, scale 1:50,000, in color.

Biswas, A., Hendry, M. J., Essilfie-Dughan, J., 2016, Geochemistry of arsenic in low sulfide-high carbonate coal waste rock, Elk Valley, British Columbia, Canada: Science of the Total Environment: <http://www.sciencedirect.com/science/article/pii/S0048969716325232>

Brown, K.G., Ross, G.L., 2002, Arsenic, Drinking Water, and Health: A Position paper of the American Council on Science and Health: Regulatory Toxicology and Pharmacology, v. 36, p. 162-174.

Casentini, B., Pettine, M., Millero, F.J., 2010, Release of arsenic from volcanic rocks through interactions with inorganic anions and organic ligands: Aquatic Geochemistry, v.16, p. 373-393.

Cunningham, W. L., Daniel, III, C. C., 2001, Investigation of Ground-Water Availability and Quality in Orange County, North Carolina: <https://nc.water.usgs.gov/reports/wri004286/pdf/report.pdf> (accessed December 2016).

Dial, A.R., Misra, S., Landing, W. M., 2015, Determination of low concentrations of iron, arsenic, selenium, cadmium, and other trace elements in natural samples using an octopole collision/reaction cell equipped quadrupole-inductively coupled plasma mass spectrometer: Rapid Communications in Mass Spectrometry, v. 29, p. 707-718.

Gray, D. B., Watts, F., Overcamp, T.J., 2001, Volatility of Arsenic in Contaminated Clay at High Temperatures: Environmental Engineering Science, v. 18, p. 1-7.

Hinkle, S. R., Polette, D. J., 1999, Arsenic in Ground Water of the Willamette Basin, Oregon: USGS, Water-Resources Investigation Report 98-4205.

Kim, D., Miranda, M. L., Tootoo, J., Bradley, P., Gelfand, A. E., 2011, Spatial Modeling for Groundwater Arsenic Levels in North Carolina: American Chemical Society, p. A-H.

McCarty, K. M., Hanh, H. T., Kim, K., 2011, Arsenic geochemistry and human health in South East Asia: Reviews on Environmental Health, v. 26, p. 71-78.

NCDEQ, n.d., Private Well Water Quality: <https://deq.nc.gov/about/divisions/water-resources/water-quality-regional-operations/groundwater-protection/ground-water-quality-monitoring/private-well-water-quality> (accessed December 2016).

Nordstrom, Kirk D., 2002, Worldwide Occurrences of Arsenic in Ground Water: Science, v. 296, p. 2143-2145.

Pippin, Charles G., 2005, Distribution of Total Arsenic in Groundwater in the North Carolina Piedmont: NWGA Naturally Occurring Contaminants Conference: Arsenic, Radium, Radon, and Uranium, p. 89-102.

Peters, Stephen C., 2008, Arsenic in groundwaters in the Northern Appalachian Mountain belt: A review of patterns and processes: Journal of Contaminant Hydrology, v. 99, pages 8-21.

Schreiber, M. E., Simo, J. A., Freiberg, P. G, 2000, Stratigraphic and geochemical controls on naturally occurring arsenic in groundwater, eastern Wisconsin, USA: Hydrogeology, v. 8, p. 161–176.

Smedley, P. L., Kinniburgh, D. G., 2002, A review of the source, behavior and distribution of arsenic in natural waters: Applied Geochemistry, v. 17, i. 5, p. 517-568.

USEPA, 2001, Technical Fact Sheet: Final Rule for Arsenic in Drinking Water: <https://nepis.epa.gov/Exe/ZyPdf.cgi?Dockey=20001XXE.txt> (accessed December 2016).

Welch, A. B., Westjohn, D. B., Helsel, D. R., Wanty, R. B., 2000, Arsenic in Groundwater of the United States: Occurrence and Geochemistry: Groundwater, v. 38, no. 4, p. 589-604.

World Health Organization, 2016, Arsenic: Fact Sheet: <http://www.who.int/mediacentre/factsheets/fs372/en/> (accessed December 2016).

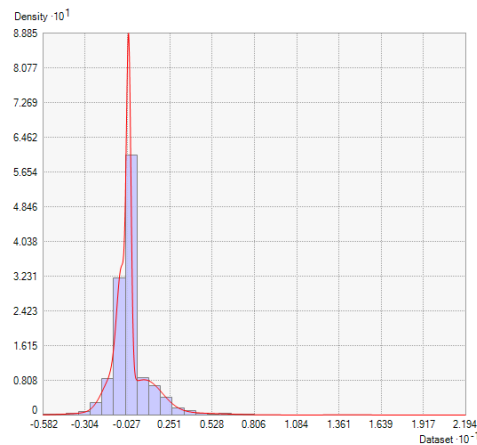
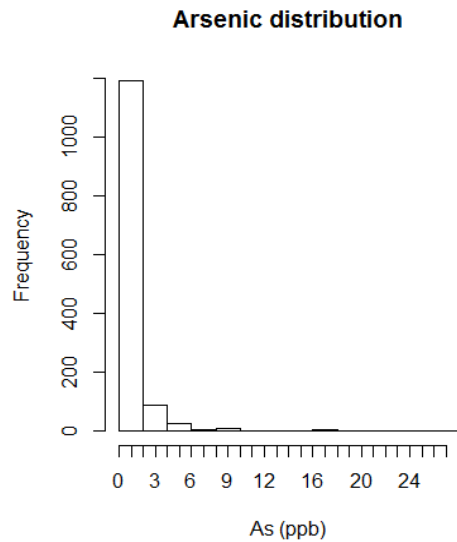
Yang, Q., Jung, H. B., Culberson, C. W., Marvinney, R. G., Loiselle, M. C., Locke, D. B., Cheek, H., Thibodeau, H., Zheng, Y, 2009, Spatial pattern of groundwater arsenic occurrence and association with bedrock geology in Greater Augusta, Maine: Environmental Science and Technology, v. 43, p. 2714–2719.

Data Sources:

Arsenic groundwater data from the North Carolina Department of Health and Human Services (NCDHHS) Laboratory.

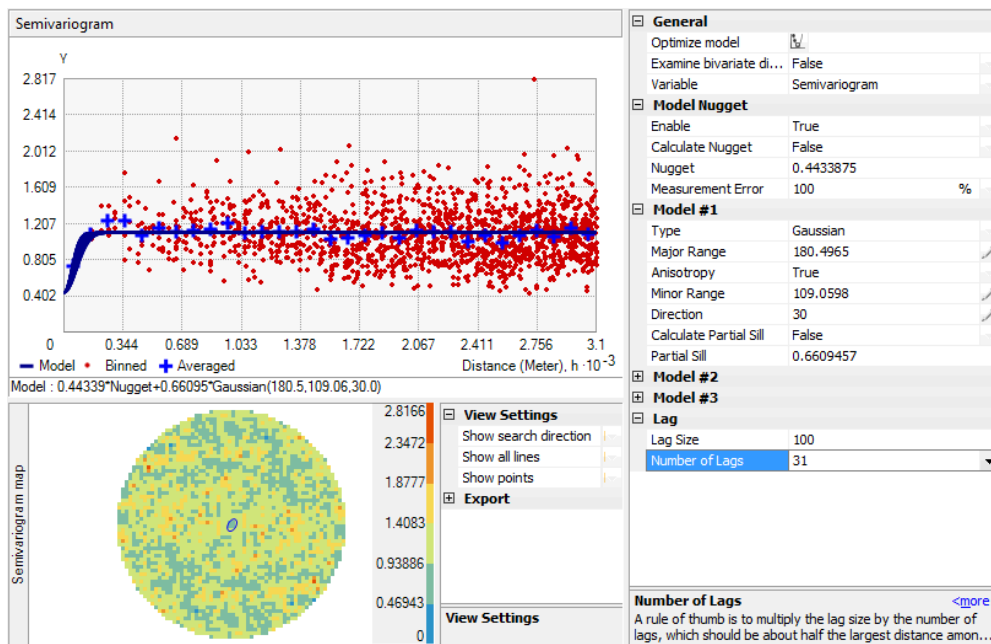
Supplementary material

Simple kriging transformation and modeling



Supp. Figure 2. Transformed data using normal score transformation

Supp. Figure. 1. Original distribution of As data



Supp. Figure 3. Simple kriging modeling parameters